

NATIONS UNIES

UNION DES NATIONS

UNITED NATIONS

COMMISSION ECONOMIQUE
POUR L'EUROPE

ECOMOMIC COMMISSION
FOR EUROPE

ECONOMIC COMMISSION
FOR EUROPE

SEMINAIRE

CONFERENCE

SEMINAR

COMMISSION DE STATISTIQUE et
COMMISSION ÉCONOMIQUE POUR
L'EUROPE



Distr.
GÉNÉRALE

CONFÉRENCE DES STATISTICIENS
EUROPÉENS

CES/SEM.43/18(Summary)
2 février 2000

Séminaire sur les systèmes intégrés
d'information statistique et les questions
connexes (ISIS 2000)
(Riga, Lettonie, 29-31 mai 2000)

FRANÇAIS
Original : ANGLAIS

Thème II : Planification et gestion des
projets statistiques

RECENSEMENTS DE L'AN 2000 : PRINCIPALES STRATÉGIES ET UTILISATION DE L'INFORMATIQUE EN SLOVÉNIE

Contribution sollicitée

de l'Office de statistique de la République de Slovénie¹

RÉSUMÉ

1. En Slovénie, cela fait plusieurs décennies que l'on fait appel, pour les recensements, aux techniques modernes de l'informatique qui ne cesse d'offrir à chaque fois de nouvelles possibilités et qui est devenue à la longue un garant d'efficacité. Ses applications se sont multipliées et s'étendent à presque toutes les étapes du recensement, depuis sa préparation jusqu'à la diffusion des données. À l'heure actuelle, les principaux problèmes se posent dans les domaines suivants :

- intégration des données provenant de différentes sources administratives et statistiques;
- réduction de la charge de travail liée aux réponses;
- intégration des opérations;
- diffusion des résultats auprès d'un public plus nombreux que jamais.

¹ Établie par Milan Katic.

2. L'Office de statistique de la République de Slovénie organisera prochainement deux grands recensements : le recensement de la population, des ménages et des habitations (prévu pour 2002) et le recensement des exploitations agricoles de l'an 2000. Pour ces deux recensements, on utilisera autant que faire se peut les données existantes stockées par diverses sources administratives. Nous disposons d'une excellente infrastructure de l'information comprenant plusieurs registres administratifs et bases de données, par exemple le Registre d'état civil, le Registre du commerce, le Registre des unités territoriales, le Registre de l'administration fiscale, le Registre de l'emploi, etc. Ces registres contiennent des identificateurs spécifiques qui permettent de relier les données avec celles d'autres registres. De la sorte, il nous est possible d'en utiliser le contenu à différentes fins, et c'est là un des objectifs stratégiques de l'Office de statistique. Considérant les bons résultats de cette démarche au cours des recensements précédents et le développement plus poussé des registres administratifs, nous avons pris la *décision stratégique* d'effectuer les futurs recensements en faisant largement appel aux registres.

3. Cette stratégie a des incidences sur le rôle et la fonction de l'informatique qui n'est plus cantonnée aux opérations techniques (transfert sur support électronique des données recueillies par questionnaires sur papier, opérations de mise en forme, présentation en tableaux, etc.), mais intervient également pour l'utilisation polyvalente et l'intégration de différentes sources de données, sous réserve toutefois d'une préparation minutieuse et d'une bonne gestion et organisation des travaux.

4. La communication décrit plus en détail les principales étapes du recensement de l'an 2000 au cours desquelles l'intervention de l'informatique est particulièrement importante, c'est-à-dire la préparation du recensement, l'entrée, la mise en forme et l'archivage des données, leur agrégation, ainsi que la production et la diffusion des résultats.

5. Les principales fonctions de l'informatique à chacune des étapes du recensement sont les suivantes :

i) Préparation du recensement

- collecte des données appropriées à caractère personnel ou autre à partir des sources existantes;
- préparation de la documentation d'appui (graphiques et cartes des districts de recensement, listes des maisons et individus);
- impression préalable sur le questionnaire des données à caractère personnel.

ii) Entrée, mise en forme et archivage des données

- intégration fonctionnelle de chacune des étapes;
- numérisation et archivage des questionnaires;
- reconnaissance optique de caractère;
- mise en forme des données avec accès direct aux questionnaires numérisés;

- conception et mise en œuvre de la base de microdonnées;
 - archivage des microdonnées.
- iii) Agrégation et analyse des données
- conception et mise en œuvre de la base des macrodonnées tirées du recensement (y compris l'agrégation à différents niveaux);
 - intégration des agrégats provenant d'autres sources;
 - soutien à des fonctions d'analyse (en particulier OLAP);
- iv) Production et diffusion des résultats
- présentation en tableaux;
 - publication des résultats (y compris sous forme électronique);
 - diffusion sur Internet et Intranet.

Préparation du recensement

6. Il faut prévoir des préparatifs à la fois vastes et minutieux. Il ne sera pas possible d'aborder les diverses tâches d'organisation et questions techniques avant d'avoir résolu les problèmes concernant le contenu et la méthode à utiliser et d'avoir arrêté les échéances (adoption d'une loi pour la circonstance).

7. L'une des tâches particulièrement importantes consiste à déterminer les sources administratives à consulter et les données à prélever auprès de ces sources. Il s'agit :

- de données à caractère personnel qui seront préalablement imprimées sur les bulletins de recensement;
- des données tirées du Registre des unités territoriales, qui sert de base pour l'établissement des cartes détaillées des districts de recensement et de la liste des logements relevant de ces districts.

8. L'établissement informatisé avant le recensement de bases de données qui fusionneront les données à caractère personnel et les données sur les districts de recensement ainsi que la préparation de la documentation mentionnée plus haut seront en partie confiés à des sous-traitants.

Entrée, mise en forme et archivage des données

9. De façon générale, l'entrée des données se fait par numérisation des questionnaires de recensement et par reconnaissance optique de caractère. Les documents (images) numérisés seront stockés dans la *base de données documentaires* et archivés. Les enregistrements obtenus par reconnaissance optique de caractère seront stockés dans la *base des données d'entrée*.

Au cours des étapes ultérieures de mise en forme des données, la base de données des documents numérisés, accessible en mode interactif servira de base pour corriger les données. Nous évaluerons les possibilités de sous-traiter la numérisation et l'entrée des données par reconnaissance optique de caractère pour le recensement de la population, car il nous reste du temps avant de devoir nous décider.

10. L'informatique intervient à toutes les étapes. *L'intégration des techniques de numérisation et de reconnaissance optique de caractère* revêt une importance particulière, et il en va de même de l'utilisation en mode interactif de la base de données documentaires au cours de l'étape de mise en forme. C'est l'Office de statistique qui effectuera la mise en forme des données (contrôle et correction) avec ses propres applications. Au cours des recensements expérimentaux, les essais de mise en forme effectués avec le système Blaise ont été couronnés de succès. Ce système sera également utilisé pour le Recensement des exploitations agricoles en juin 2000.

Agrégation et analyse

11. On procèdera à l'agrégation à partir des données vérifiées contenues dans la base des microdonnées, et on établira ainsi la base des macrodonnées. Les données seront agrégées au niveau du district statistique, considéré comme l'unité territoriale qui permet un regroupement en unités territoriales plus importantes, par exemple de grands ensembles, des municipalités, des régions, voire le pays tout entier.

12. La base des macrodonnées contiendra, en plus des agrégats du recensement, d'autres données agrégées provenant de sources administratives ou statistiques (les recensements de 1998 et de 1991 par exemple). Ainsi donc, il sera créé une seule base de données pour diverses analyses ainsi que pour des travaux de recherche sur des phénomènes socioéconomiques à réaliser après le recensement.

13. À ce stade, l'informatique sert à intégrer :

- intégration d'agrégats du recensement et de données provenant d'autres sources;
- facilitation de fonctions d'analyse en ligne avec les outils correspondants de l'OLAP.

Production et diffusion des résultats

14. C'est pour cette étape que le rôle de l'informatique a le plus évolué par rapport aux recensements précédents. Les résultats du recensement seront disponibles sur l'Internet et diffusés non seulement sur papier mais également sous forme électronique, ce qui ouvre un éventail de possibilités d'utilisation à grande échelle des données à des fins d'analyse ou de traitement. Il est prévu de mettre en place un système de présentation en tableaux, qui produira tous les tableaux reprenant les résultats du recensement et permettra aussi de publier et diffuser les résultats du recensement à la fois sur papier et sur support électronique. Il est également prévu d'établir un atlas de la Slovénie dans lequel les principales données du recensement seront présentées sous forme de cartes et de graphiques. Les agrégats au niveau le plus élevé seront accessibles sur l'Internet.
